

次回アクセスを保証する Web システム「NAP-Web」の開発

05G455 加地 智彦（最所研究室）

Web サーバが過負荷状態に陥った際に、アクセス制御により負荷を下げるが、この際にアクセスを拒絶されたユーザに対して次回アクセスを保証する Web システム「NAP-Web」を開発している。本稿では当システムの概要と基盤モジュールであるアクセススケジューリング機構について述べる。

1 はじめに

近年のインターネットの普及により、Web サーバの負荷は増大し続けている。負荷の増大に対応するために、ミラーリングなどのサーバ負荷分散技術が発展してきた。しかしながら、サービス運営者が予測したものより大きな処理が求められた場合、Web サーバが過負荷になることは依然として避けられない。

本研究では Web サーバが過負荷になったときに生じるユーザの不満、すなわち「いつになればサービスを受けることができるのか」という点に着目し、この不満を緩和する Web システム「NAP-Web」を提案している [1]。このシステムでは、システムが過負荷にならないように同時処理可能なアクセス数を制限するが、これによってアクセスを拒絶されたユーザに対していつになればアクセスが可能になるかを保証することでユーザの不満を軽減する。

本稿では、「NAP-Web」の概要を述べた後、本システムの基盤モジュールとして作成したアクセススケジューリング機構の概要と評価した結果について述べる。

2 「NAP-Web」の概要

ニュースに取り上げられるなどの外的要因により、Web サーバのトラフィックが急激に増加することがある。増大したトラフィックにより過負荷状態に陥った Web サーバはサービスを提供できなくなる。急激なトラフィックの増加を予測するのは困難であるし、増大した負荷に耐え切れるだけの設備を常時準備するのはコストの面から現実的ではない。

本研究では、サーバが過負荷状態になり、サービスを受けることが出来なくなったユーザの不満を緩和することを目的とした Web システム「Next Access Promissed Web(NAP-Web)」を提案している。NAP-Web はサーバが過負荷状態に陥った際にユーザに対して「次、いつアクセスすればサービスの提供が受けられるのか」を保証することでユーザの不満を緩和する。本システムは Web サーバデーモン「Apache」を基盤として開発している。

NAP-Web は以下の三つの機能を持つ。

- 動的負荷制御機能
- アクセス拒絶されたユーザに対して次回アクセスを保証する機能
- クライアント・サーバ連携による次回アクセス補助機能

3 アクセススケジューリング機構の概要

動的負荷制御機能とアクセス拒絶されたユーザに対して次回アクセスを保証する機能を実現するために、アクセススケジューリング機構を試作した。本機構は 1) クライアントへの応答時間を知ること、2) 応答時間の予測の二つの要件を満たすことを目的としている。

現在最も利用されている Web サーバデーモン Apache では、同時処理の最大数を制限している。この制限を越えたアクセスが同時に来た場合は Accept 待ちになり、クライアントへの応答時間をサーバ側で知ることが出来ない。また、同時処理アクセスが多くなると、応答時間のばらつきが大きくなり、応答時間の予測が難しくなることを確認している。[1]

このため、実際に処理するアクセスの数を制限し、制限を越えたアクセスをキューイングする機構を開発した。本機構は、クライアントからのアクセスを全て Accept する。そしてこれらのアクセスを Figure1 に示すように三つのグループに分けてスケジューリングを行う。「Run_Ready」、「Wait」、「Next_Wait」はそれぞれ、実際に処理するアクセス、キューイングされるアクセス、拒絶され再度アクセスすることが予想されるアクセスのグループである。

現在のところ、「Run_Ready」アクセスグループと「Wait」アクセスグループを扱う部分のモジュールの実装を行った。「Next_Wait」アクセスグループについては未実装である。「Next_Wait」アクセスグループに該当するアクセスに関しては HTTP エラー 503 のステータスによりアクセス拒否を行っている。

4 アクセススケジューリング機構の評価

前節で提案したアクセススケジューリング機構について 1、2 の要件を満たすことが出来るか評価を行った。まず、要件 1 についての調査実験を行った。掲

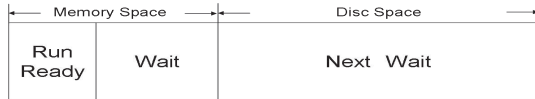


Figure 1: 3つのアクセスグループ

示板プログラムに対して、150個のスレッドを生成し、10KBの文字データをPOSTするアクセス要求を60秒間で60回繰り返す負荷テストを機構の適用前後で行った。サーバ側の設定は機構の適用前はMaxClientsを10とした。適用後はMaxClientsを1,000とし、スケジューリング機構用の設定としてRun_Readyアクセスグループの最大値を10、Waitキューの大きさを200とした。機構の適用前後のクライアント及びサーバでの応答時間のグラフをFigure2, Figure3に示す。機構の適用前後でクライアントとサーバで計測した応答時間の相関に大きな差があることが分かる。適用前の相関係数は0.62、適用後の相関係数は0.99であり、機構を適用することによって要件1を満たすことが可能であると分かる。

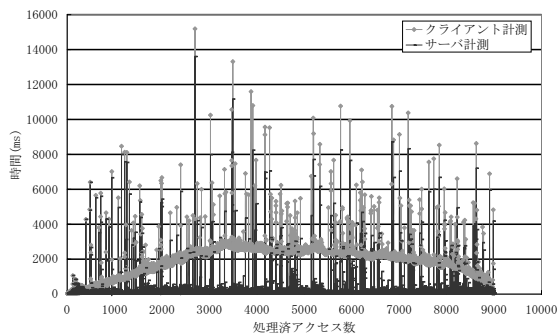


Figure 2: 機構適用前の負荷実験

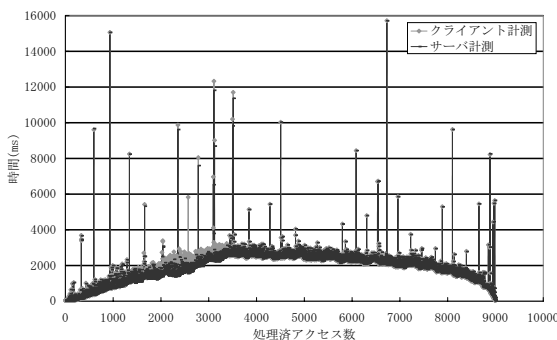


Figure 3: 機構の適用後の負荷実験

次に、応答時間の予測が可能かどうかの点についての実験を行った。Run_Readyグループの数を10個、Waitキューの大きさを200に固定し、機構を適用した

Webサーバに対して1KBの文字データをPOSTする内容のアクセスを同時に50、100、150個と変え、それぞれ300秒間行う負荷テストを行った。Figure4はこの負荷テストの結果において、Waitキューに入っているアクセスの数と応答時間の関係を示すグラフである。ある程度のばらつきはあるものの、Waitキューの中のアクセスの数と応答時間がほぼ正比例している。アクセススケジューリング機構を組み込むことによって、応答時間の予測が可能であることが分かる。

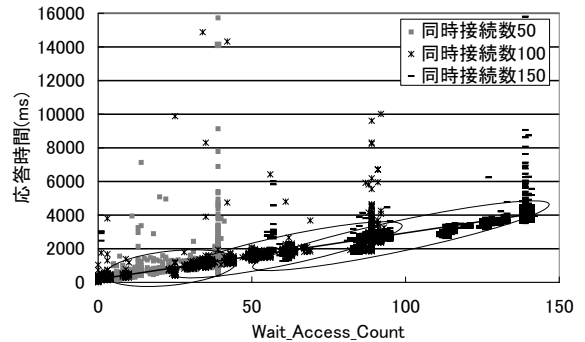


Figure 4: Waitアクセスキューの大きさと応答時間の関係

5 まとめ

Webサーバの過負荷時のユーザの不満を緩和するためのWebシステム「NAP-Web」を提案し、このシステムの基盤モジュールとなるアクセススケジューリング機構を試作した。この機構はWebアクセスをグループ分けしてスケジューリングを行う。この機構を用いることで、クライアントから見える応答時間をサーバ側から計測することが可能になった。そして、応答時間の予測が可能になった。

今後の課題として、

- 応答時間予測手法の確立
- 適切な閾値の動的制御方法の確立

などが挙げられる。

References

- [1] 加地智彦、最所圭三，“Web処理を効率化するアクセススケジューリング機構について”，情報処理学会研究報告、2007-OS-104、pp.75-80、2007.